

University of Groningen

Incomplete contour representations and shape descriptors

Ghosh, Anarta; Petkov, Nicolai

Published in:
BRAIN, VISION, AND ARTIFICIAL INTELLIGENCE, PROCEEDINGS

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2005

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Ghosh, A., & Petkov, N. (2005). Incomplete contour representations and shape descriptors: ICR test studies. In MD. Gregorio, DiMaio, M. Frucci, & C. Musio (Eds.), *BRAIN, VISION, AND ARTIFICIAL INTELLIGENCE, PROCEEDINGS* (pp. 416-425). (LECTURE NOTES IN COMPUTER SCIENCE; Vol. 3704). Springer.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Incomplete Contour Representations and Shape Descriptors: ICR Test Studies

Anarta Ghosh and Nicolai Petkov

Institute of Mathematics and Computing Science, University of Groningen,
P.O.Box. 800, 9700 AV Groningen, The Netherlands
{anarta, petkov}@cs.rug.nl

Abstract. Inspired by psychophysical studies of the human cognitive abilities we propose a novel aspect and a method for performance evaluation of contour based shape recognition algorithms regarding their robustness to incompleteness of contours. We use complete contour representations of objects as a reference (training) set. Incomplete contour representations of the same objects are used as a test set. The performance of an algorithm is reported using the recognition rate as a function of the percentage of contour retained. We call this evaluation procedure the ICR test. We consider three types of contour incompleteness, viz. segment-wise contour deletion, occlusion and random pixel depletion. We illustrate the test procedure using two shape recognition algorithms. These algorithms use a shape context and a distance multiset as local shape descriptors. Both algorithms qualitatively mimic human visual perception in the sense that the recognition performance monotonously increases with the degree of completeness and that they perform best in the case of random depletion and worst in the case of occluded contours. The distance multiset method performs better than the shape context method in this evaluation framework.

1 Introduction

We can easily recognize the butterflies depicted in Fig. 1, even though 50% of the contour is removed segment-wise in the left image, the right half of the contour is not visible in the middle image, and 80% of the contour points have been removed (randomly) in the right image. Psychologist E. S. Gollin [6] investigated this human ability to recognize objects from incomplete contour representations. The main objective of his study was to investigate the performance of humans in recognizing objects with incomplete contours as a function of developmental characteristics, such as mental and chronological age and intelligence quotient. As subjects of his experiments he chose children of different age groups and a group of adults. In his experiments Gollin used sets of contour images with different degrees of incompleteness (Fig. 2) and addressed the following questions: (1) In order to be recognized, how complete the contours of common objects need to be? (2) How does training affect the recognition performance in case of incomplete representations? Through his experiments he found that human

ability to recognize objects with incomplete contours (a) depends on intelligence quotient and (b) is improved by training.

This aspect of recognition of objects with incomplete contours is also very important in the context of processing visual information using computers. A natural image and two edge images, obtained from it are shown in Fig. 3. The middle image was obtained by applying a bank of Gabor energy filters [8]. It contains the contours of the object of interest, viz. a gazelle, but it also contains a large number of texture edges in the background that are not related in any way to the shape of the gazelle. There would be a devastating effect of these texture edges on any currently known contour based shape recognition algorithm. Advanced contour detection methods based on surround suppression [8,9] succeed in separating the essential object contours from the texture edges, as illustrated

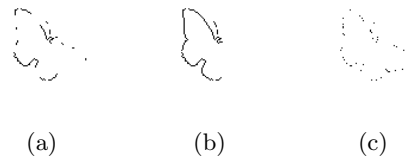


Fig. 1. A butterfly can be recognized even though (a) 50% of its contour has been removed segment-wise, (b) one of its wings is not fully visible (occluded), (c) 80% of the contour pixels have been randomly removed

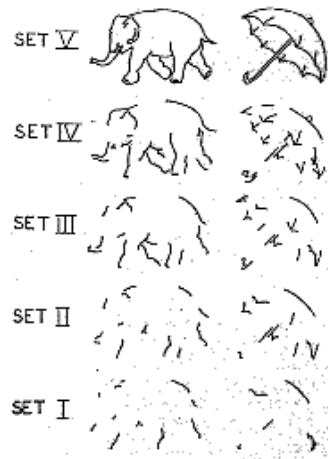


Fig. 2. Example of image sets used in Gollin's original test [1]. The images in set V are complete contour representations and the other sets are derived from set V by removing segment-wise an increasing fraction of the contour. Reproduced with the permission from the author and the publisher of: E. S. Gollin, Developmental studies of visual recognition of incomplete objects. *Perceptual and Motor Skills*, Vol. 11 pp. 289-298, 1960, copyright Southern University Press.

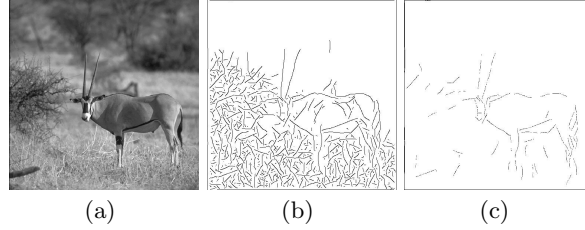


Fig. 3. (a) Image of a gazelle in its natural habitat. (b) Result of edge detection with a bank of Gabor energy filters. (c) Result of contour detection by a bank of Gabor energy filters augmented with a biologically motivated surround suppression of texture edges.

by the right-most image in Fig. 3, but at the same time these methods have a certain negative side effect of depleting the contours of the objects of interest. Hence in practical situations the robustness of shape recognition methods to contour incompleteness is also an issue of importance.

Inspired by Gollin’s study we put forward a novel attribute, viz. *robustness to incomplete contour representations*, that any contour based object recognition system/algorithm should have. We choose an idealized situation where: (a) complete contour representations of the objects to be recognized form the reference (training) set or “memory” of the system/algorithm, (b) incomplete contour representations of the same objects are derived from the afore mentioned complete representations and are used as a test set, (c) the performance of the system/algorithm in recognizing the objects from these incomplete representations is evaluated. The main reason behind evaluating the performance of object recognition algorithms in such an ideal situation is the rational logic that in order to perform well in a real world scenario (natural images) any recognition system should first perform well in such idealized (simple) situations.

We investigate the performance of two contour based shape recognition methods, which use a shape context [1] and a distance multiset [7] as shape descriptors, by comparing an object represented by incomplete contours with all objects in a reference set represented by complete contours and determining the nearest neighbor. If the nearest neighbor is the object from which the incomplete contour representation is derived we consider the recognition to be correct, otherwise incorrect. As incomplete contour representations of an object, in addition to Gollin’s method of segment-wise contour deletion (Fig. 2) we also consider other types of incompleteness, viz. occlusion and random pixel depletion. We name the corresponding studies segment-wise deletion test, occlusion test, and depletion test and collectively call these tests in short *Incomplete Contour Representation (ICR) tests*.

In Section 2 we describe the shape recognition methods which we use for illustration. The experimental design and the achieved results are discussed in Section 3. A summary and conclusions are presented in Section 4.

2 Shape Recognition Methods

The recognition of objects in the methods studied below is done by computing dissimilarity between the contour representations of two objects by using a point correspondence paradigm. Shape descriptors associated with the points are used to find the point correspondences. To maintain brevity and focus on the illustration of the ICR test framework we use simpler versions of the algorithms described in [7] and [1].

2.1 Distance Multiset

The *distance multiset* for a point p in the contour of an object \mathcal{O} of N points, is formally defined as the following vector [7]: $D_N^{\mathcal{O}}(p) = (\ln(d_1(p)), \ln(d_2(p)), \dots, \ln(d_{N-1}(p)))$ where $d_j(p)$ is the Euclidean distance between p and its j^{th} nearest neighbor in \mathcal{O} . In this approach the shape of an object $\mathcal{O} \equiv \{p_1 \dots p_N\}$ defined by a set of contour points is described by the set of distance multisets in the following way: $S_{\mathcal{O}}^{DM} \equiv \{D_N^{\mathcal{O}}(p) | p \in \mathcal{O}\}$. Next, a cost $c(X, Y)$ of matching two distance multisets X and Y is defined and computed by using the algorithm described in [13]. Let $c_{i,j}^{DM}$ be the cost of matching a point p_i in an object \mathcal{O}_1 represented by M contour points to a point q_j in an object \mathcal{O}_2 represented by N contour points, $M \leq N$: $c_{i,j}^{DM} \equiv c(D_N^{\mathcal{O}_1}(p_i), D_M^{\mathcal{O}_2}(q_j))$. Then the dissimilarity between the shapes $S_{\mathcal{O}_1}^{DM}$ and $S_{\mathcal{O}_2}^{DM}$ is defined as follows: $d^{DM}(S_{\mathcal{O}_1}^{DM}, S_{\mathcal{O}_2}^{DM}) \equiv \sum_{i=1}^M \min\{c_{i,j}^{DM} | j = 1 \dots N\}$.

2.2 Shape Context

The *shape context* [1] of a point p belonging to the contour of an object is a bi-variate histogram in a log-polar coordinate system that gives the distribution of contour points in the surroundings of p . Let an object \mathcal{O} be represented by a set of contour points, $\mathcal{O} \equiv \{p_1 \dots p_N\}$. Formally, the authors of this method define the shape context of a point $p \in \mathcal{O}$ as a vector in the following way: $H_K^{\mathcal{O}}(p) = (h_1(p), h_2(p), \dots, h_K(p))$, where $h_k(p) = \text{card}\{q \neq p | q \in \mathcal{O}, (q - p) \in \text{bin}(k)\}$ is the number of contour points in the k^{th} bin $\text{bin}(k)$ and K is the total number of histogram bins. The bins are constructed by dividing the image plane into K partitions (in a log-polar coordinate system) with p as the origin. In this study we use 5 intervals for the log distance r , and 12 intervals for the polar angle θ , so $K = 60$. As suggested in [1], we randomly choose 100 points (if available) from the contour of an object and calculate their shape contexts. The shape of the object is described using the set of shape contexts associated with the contour points in the following way: $S_{\mathcal{O}}^{SC} \equiv \{H_K^{\mathcal{O}}(p) | p \in \mathcal{O}\}$. The cost of matching a point p_i that belongs to the contour of an object \mathcal{O}_1 of M points, to a point q_j from the contour of an object \mathcal{O}_2 of N points is defined as follows: $c_{i,j}^{SC} \equiv \frac{1}{2} \sum_{k=1}^K \frac{[h_k(p_i) - h_k(q_j)]^2}{h_k(p_i) + h_k(q_j)}$, which yields an $M \times N$ cost matrix of point-wise dissimilarities. Next we compute the dissimilarity between the shapes $S_{\mathcal{O}_1}^{SC}$ and $S_{\mathcal{O}_2}^{SC}$ of the objects in the following way: $d^{SC}(S_{\mathcal{O}_1}^{SC}, S_{\mathcal{O}_2}^{SC}) \equiv \sum_{i=1}^M \min\{c_{i,j}^{SC} | j = 1, \dots, N\}$.

3 Experiments and Results

3.1 Image Set

We choose the silhouette images from the MPEG-7 database [10] as our dataset. In this dataset there are 1400 images divided into 70 classes, each of 20 similar objects (eg. apple, bird, bat, etc). One object from each class is chosen and the contours of the objects are extracted using Gabor filters [8]. These 70 contour images are rescaled in such a way that the diameter (maximum Euclidean distance between contour pixels) is approximately the same (76 pixels) for all objects (Fig. 4). These 70 rescaled contour images are used as the reference ("memory" of the recognition system) images in our experiments. These images are analogous to the complete representations, set V of Fig. 2, used in Gollin's original study.

Incomplete contour representations of objects for the *segment-wise deletion test* are constructed by randomly removing continuous segments of the contours and retaining a given percentage of contour pixels from the above mentioned complete contour representations. For c percent of retained pixels approximately $\lceil \log_2(\frac{100-c}{8}) \rceil$ segments are deleted. Incomplete representations for the *occlusion test* are created by removing a given percentage of consecutive contour pixels starting from the leftmost (Fig. 5(b)) or the rightmost pixel (Fig. 5(c)) of an object. The left and right occlusion are deliberately chosen due to the fact that in case of natural images the object of interest is most commonly occluded either from the left or from the right. A random pixel deletion is performed to construct the incomplete representations for the *depletion test* (Fig. 5(d)). The percentages of retained pixels are chosen in the following way: from 2% to 4% in steps of 1%, from 5% to 85% in steps of 5%, and 100% for the depletion test; from 5% to 85% in steps of 5%, and 100% for the segment-wise deletion and the occlusion tests. For each type (segment-wise deletion, occlusion and depletion) and degree of contour degradation 70 test images are created. In the web-site www.cs.rug.nl/~petkov the complete dataset for the proposed ICR test is available.

3.2 Methodology

A test image (incomplete contour representation of an object) obtained from one of the 70 reference images is compared with all 70 reference images using a given

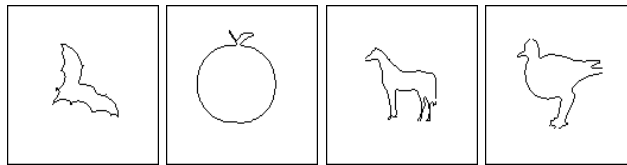


Fig. 4. Rescaled contour images obtained from samples of MPEG-7 silhouette database. These images are considered as complete representations that comprise the memory of the recognition system.

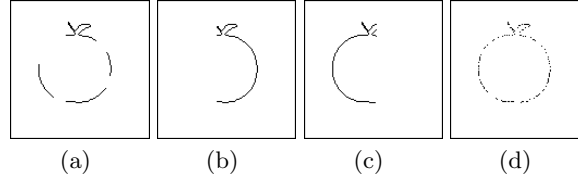


Fig. 5. Incomplete contour representations: (a) Segment-wise deleted contour representation (this type of incompleteness corresponds to incomplete representations of Gollin’s original study, set I to IV of Fig. 2). (b) Left-occluded contour representation. (c) Right-occluded contour representation. (d) Depleted contour representation.

shape comparison algorithm, described in Section 2 and a decision is taken about which reference image the degraded image is most similar to (nearest neighbor search). If the nearest neighbor is the reference image from which the degraded image was obtained, the recognition is considered correct, otherwise incorrect. If the nearest neighbor is found to be not unique then the recognition is also considered incorrect. For each of the three tests (segment-wise deletion, occlusion, depletion) and for each degree (c percentage of retained contour pixels) of contour image degradation, the corresponding 70 test images are compared with each of the 70 reference images and the percentage of correct recognition $P(c)$ is determined. An average of the recognition rates with left and right occluded images for a given percentage of retained contour is computed to evaluate the performance of the algorithms in the occlusion test.

3.3 Results and Discussions

The results of our experiments are illustrated in Fig. 6. The recognition rate is a monotonic increasing function of the percentage of contour retainment in all three tests. In this respect the considered algorithms resemble the human visual system [3,4,14]. Both methods perform worst in the occlusion test and best in the depletion test, which also conforms with the recognition performance of humans, as occluded contour images carry the least amount of shape information and depleted contour images carry maximum shape information in the context of human visual perception.

The performance of the distance multiset method is appreciably better than that of the shape context method for any percentage of retained contour pixels in the case of the segment-wise deletion test (Fig. 6(top left)) and the occlusion test (Fig. 6(top right)). From the results of the depletion test (Fig. 6(bottom)) we see that both the shape context method and the distance multiset method perform very well in recognizing objects with depleted contour representations, if more than 40% and 5%, respectively, of the contour points are retained. For higher degree of depletion ($c \leq 40$) the distance multiset method outperforms the shape context method.

The better performance of the distance multiset method in general can be explained by the fact that the proposed ICR tests give advantage to the algorithms

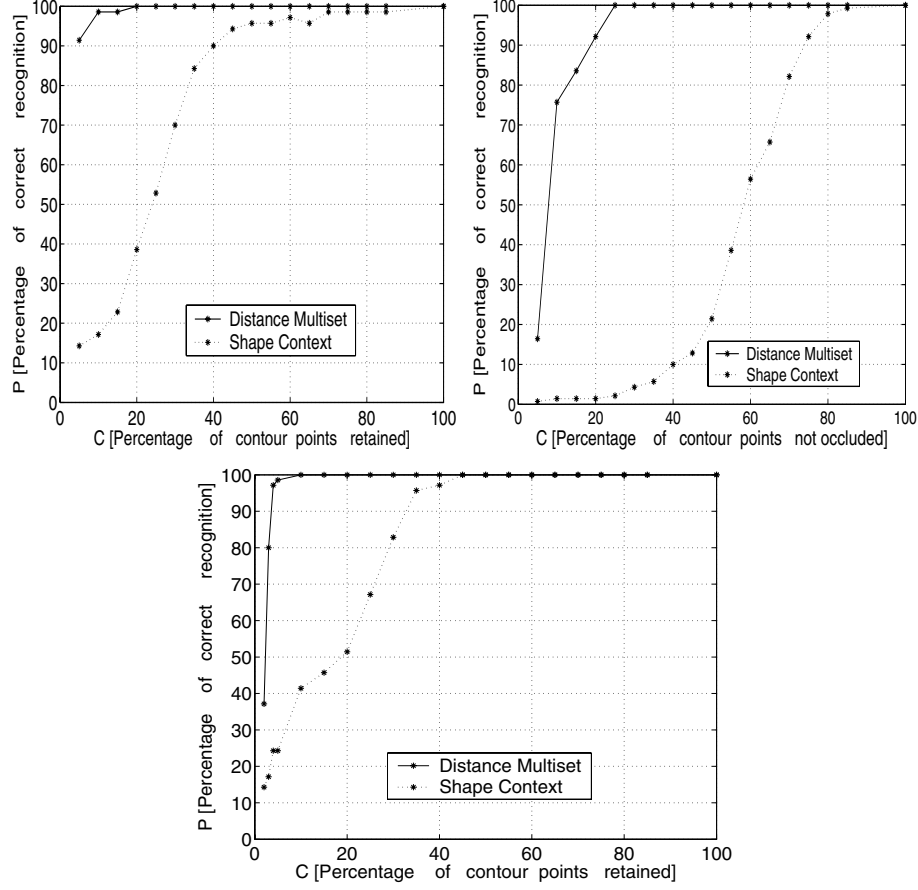


Fig. 6. Results of the ICR tests with a subset of MPEG-7 dataset: (top left) Segment-wise deletion test; (top right) Occlusion test; (bottom) Depletion test

which yield zero dissimilarity in a comparison of two objects represented by two sets of points where one is a subset of the other. This property of the distance multiset algorithm is explained in more detail below. Let two sets $A, B, \subset \mathbf{R}^2$ be such that $B = \{f(\mathbf{x}) : \mathbf{x} \in A\}$, where $f : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ is an isometry.

Lemma 1. *If $B = f(A)$, $C \subset B$ and $\text{card}(C) \geq 2$. then $d^{DM}(S_C^{DM}, S_A^{DM}) = 0$ where S_C^{DM} and S_A^{DM} are the shapes, described by distance multisets, corresponding to C and A , respectively.*

For the proof of the lemma refer to [5]. In our study A corresponds to the set of contour points of a reference object, f is the identity transformation (i.e. $B = A$) and C is the set of contour points of an incomplete representation.

The implication of the lemma is two-fold: (1) The recognition will be incorrect by the distance multiset method only when the nearest neighbor of a test object

in the reference set is not unique. (2) The distance multiset method should perform exactly the same way when f is not the identity transformation [5].

The above lemma does not hold for the shape context method, but this method can be modified in such a way that the relation $d^{SC}(S_{\mathcal{O}}^{SC}, S_{\mathcal{O}'}^{SC}) = 0$ can be approximately fulfilled if \mathcal{O}' is an incomplete representation (of modest degree) derived from \mathcal{O} . Specifically, we normalize the shape context $H_K^{\mathcal{O}}(p)$ by dividing its elements by the total number of points $\text{card}(\mathcal{O})$ in the corresponding object \mathcal{O} . If $\mathcal{O}' \subset \mathcal{O}$ is an incomplete representation derived from \mathcal{O} and if $H_K^{\mathcal{O}'}(p)$ is the normalized (by $\text{card}(\mathcal{O}')$) shape context of a point p ($p \in \mathcal{O}'$) in this incomplete representation, the relation $H_K^{\mathcal{O}'}(p) \approx H_K^{\mathcal{O}}(p)$ will hold for modest degrees of contour deletion because the ratio of the number of contour points in each bin to the total number of points will be approximately the same for the complete and the incomplete contour representations and hence we have $d^{SC}(S_{\mathcal{O}}^{SC}, S_{\mathcal{O}'}^{SC}) \approx 0$ for normalized shape context. In our experiments we found that the performance of the shape context method is greatly improved due to this normalization in the segment-wise deletion and the depletion test [5].

As the scope of this paper is to introduce a new test, it is important to check if the conclusions drawn from the ICR test are consistent across *datasets*. We carried out experiments using a second set of images, the Columbia University Image Library (COIL-20) dataset and (compared to the MPEG-7 dataset) no qualitative difference in the performances of the algorithms was observed, c.f [5].

The *object size* can have effect on the results of an ICR test through (a) the resolution of the reference objects and (b) a possible mismatch between the size of a reference object and a test object. Regarding the resolution of the reference objects, in our experiments we found that for a given percentage of contour degradation (by any method) the performance of the algorithms grows with the diameter of the reference objects. To eliminate this effect and to standardize the test procedure we rescaled the reference contour images to a fixed diameter (76 pixel units). For more detailed discussion on the effect of the object size on the proposed ICR test refer to [5]. The problem of a possible mismatch between the sizes of reference and test objects can be dealt with either by using scale invariance procedures prescribed in [1], [7] or by using a multiscale approach.

The performance curves obtained in the ICR tests can be used to compare algorithms as illustrated in Fig. 6. To define a *criterion for acceptable performance* of an algorithm in the ICR test, the performance of humans in similar experimental [3,4,14] setup can be used as a reference [5].

A good performance in the original ICR test does not guarantee good performance in other respects, e.g. robustness to shape or size variation. Hence, a good performance in the ICR test should be considered as a *necessary condition* for object recognition methods to perform well in a real world scenario but not as a sufficient one. We are not aware of any evaluation procedure for shape recognition methods which is sufficient in such respect. The basic framework of the ICR test proposed in this paper can easily be extended to test robustness of algorithms to more than one criterion, e.g. a *bull's eye* ICR test for evaluat-

ing robustness to shape variations along with robustness to incomplete contour representations. We present results of such a bull's eye ICR test in [5].

4 Summary and Conclusion

Shape descriptor based object recognition methods have been evaluated and compared using various characteristics like invariance, uniqueness and stability [12]. Marr and Nishihara [11] proposed three criteria for judging the effectiveness of a shape descriptor, viz. accessibility, scope and uniqueness, stability and sensitivity. Brady [2] put forward a set of criteria for representation of shape, viz. rich local support, smooth extension and propagation. In the current work, motivated by characteristics of the human visual system [6], we propose an additional new criterion, viz. robustness to contour incompleteness to compare and characterize contour based shape recognition algorithms using their performance in recognizing objects with incomplete contours. We are not aware of any such comparison and characterization in the present literature.

We put forward the following procedure which we call the ICR test: (1) Take a set of images of objects and extract contours. Rescale all contour images to the same object diameter. (2) Train the recognition system with these complete contour representations. (3) Construct different sets of incomplete representations from the complete contour representations; quantify the level of incompleteness using the percentage of contour pixels retained. (4) Using the incomplete representations as a test set evaluate the recognition rate as a function of the percentage of contour pixels retained.

To illustrate the framework we use two shape recognition methods based on the shape context and the distance multiset. The two methods tested were chosen merely for illustrative purposes and we did not aim to prove superiority of any method. A complete comparative study of the two methods is out of the scope of this work. In our illustrative experiments we found that: (1) The distance multiset shape recognition method outperforms the shape context method. (2) Both methods perform similar to the human visual system in the sense that their performances are increasing functions of the degree of contour completeness and are best in the case of the depletion test and worst in the case of the occlusion test.

References

1. S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 24, pages 509–522, 2002.
2. M. Brady. Criteria for representations of shape. In J. Beck, B. Hope and A. Rosenfeld, editors, *Human and Machine Vision*, Academic Press, pages 39–84, 1983.
3. V. Chihman, V. Bondarko, Y. Shelpin, and M. Danilova. Fragmental figure perception. *Perception*. Vol. 33 Supplement, page 76a, 2004.

4. N. P. Foreman and R. Hemmings. The Gollin incomplete figures test: A flexible, computerised version. *Perception. Vol. 16*, pages 543–548, 1987.
5. A. Ghosh and N. Petkov. Robustness of shape descriptors to incomplete contour representations. *IEEE Transactions on Pattern Analysis and Machine Intelligence, In press*, 2005.
6. E.S. Gollin. Developmental studies of visual recognition of incomplete objects. *Perceptual and Motor Skills. Vol. 11*, pages 289–298, 1960.
7. C. Grigorescu and N. Petkov. Distance sets for shape filters and shape recognition. *IEEE Transactions on Image Processing, Vol. 12, No. 10*, pages 1274–1286, 2003.
8. C. Grigorescu, N. Petkov, and M. Westenberg. Contour detection based on non-classical receptive field inhibition. *IEEE Transactions on Image Processing, Vol. 12, July*, pages 729–739, 2003.
9. C. Grigorescu, N. Petkov, and M. A. Westenberg. Contour and boundary detection improved by surround suppression of texture edges. *Image Vision and Computing, Vol. 22*, pages 609–622, 2004.
10. L.J. Latecki, R. Lakämper, and U. Eckhardt. Shape descriptors for non-rigid shapes with single closed contour. *In Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, pages 424–429, 1998.
11. D. Marr and H.K. Nishihara. Representation and recognition of the spatial organization of three dimensional shapes. *In Proc. Roy. Soc. London, B., Vol. 200*, pages 269–294, 1978.
12. F. Mokhtarian and A.K. Mackworth. A theory of multiscale, curvature-based shape representation for planar curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, No. 8*, pages 789–805, 1992.
13. N. Petkov. Algorithm for the cost of an optimal assignment of two sets of real numbers. *Technical report, 2003-9-07, Institute of Mathematics and Computing Science, University of Groningen*, 2003.
14. Y. Shelepin, O. Vahromeeva, A. Harauzov, S. Pronin, N. Foreman, and V. Chihman. Recognition of incomplete contour and half-tone figures. *Perception. Vol. 33 Supplement*, page 85c, 2004.